

Prediction Probabilities From Foreshocks

DUNCAN CARR AGNEW

*Institute of Geophysics and Planetary Physics,
University of California, La Jolla*

LUCILE M. JONES

U.S. Geological Survey, Pasadena, California

When any earthquake occurs, the possibility that it might be a foreshock increases the probability that a larger earthquake will occur nearby within the next few days. Clearly, the probability of a very large earthquake ought to be higher if the candidate foreshock were on or near a fault capable of producing that very large mainshock, especially if the fault is towards the end of its seismic cycle. We derive an expression for the probability of a major earthquake characteristic to a particular fault segment, given the occurrence of a potential foreshock near the fault. To evaluate this expression, we need: (1) the rate of background seismic activity in the area, (2) the long-term probability of a large earthquake on the fault, and (3) the rate at which foreshocks precede large earthquakes, as a function of time, magnitude, and spatial location. For this last function we assume the average properties of foreshocks to moderate earthquakes in California: (1) the rate of mainshock occurrence after foreshocks decays roughly as t^{-1} , so that most foreshocks are within three days of their mainshock, (2) foreshocks and mainshocks occur within 10 km of each other, and (3) the fraction of mainshocks with foreshocks increases linearly as the magnitude threshold for foreshocks decreases, with 50% of the mainshocks having foreshocks with magnitudes within three units of the mainshock magnitude (within three days). We apply our results to the San Andreas, Hayward, San Jacinto, and Imperial faults, using the probabilities of large earthquakes from the report of the Working Group on California Earthquake Probabilities (1988). The magnitude of candidate event required to produce a 1% probability of a large earthquake on the San Andreas fault within three days ranges from a high of 5.3 for the segment in San Geronio Pass to a low of 3.6 for the Carrizo Plain.

Probably the most evil feature of an earthquake is its suddenness. It is true that in the vast majority of cases a severe shock is heralded by a series of preliminary shocks of slight intensity [but] only after the havoc has been wrought does the memory recall the sinister warnings of hypogene action.

C. G. Knott (1908, p. 10)

1. INTRODUCTION

Many damaging earthquakes have been preceded by smaller earthquakes that occur within a few days and a few kilometers of the mainshock [e.g., Jones and Molnar, 1979]; these are referred to as immediate foreshocks. If such foreshocks could be recognized before the mainshock, they would be very effective for short-term earthquake prediction; but so far no way has been found to distinguish them from other earthquakes. Even without this, the mere existence of foreshocks provides some useful predictive capacity. When any earthquake occurs, the possibility that it might be a foreshock increases the probability that a larger earthquake will soon happen nearby. For southern California, Jones [1985] showed that after any earthquake there is a 6% probability that a second one equal to or larger than the first will follow within five days and 10 km of the first. The probability is much lower for a second earthquake much larger than the first; for example, the probability of an earthquake two units of magnitude larger is only 0.2%. Using

these results, the U.S. Geological Survey has issued four short-term earthquake advisories after moderate earthquakes [e.g., Goltz, 1985]. A more recent study by Kagan and Knopoff [1987] developed a model for the clustering of earthquakes which could indicate areas of space and time in which larger events might follow smaller ones. The size of these areas depended on the probability gain, the ratio of probability of an earthquake given the occurrence of a possible precursor (such as a foreshock) to the probability in the absence of such a precursor [Kagan and Knopoff, 1977; Vere-Jones, 1978; Aki, 1981]. For low levels of probability gain, Kagan and Knopoff [1987] found that one-third of all earthquakes with magnitudes 4 and above fell within their predicted regions.

These results are from studies of earthquake catalogs; Jones [1985] used a catalog for southern California, and Kagan and Knopoff [1987] used one for central California. As a consequence, both papers give generic results about pairs of earthquakes, without much regard for other factors. But it ought to be possible to do better: the probability of a very large earthquake should be higher if the candidate foreshock were to occur near a fault capable of producing that mainshock than if it were located in an area where we believe such a mainshock to be very unlikely. Moreover, the chance of a candidate earthquake actually being a foreshock should be higher if the rate of background (nonforeshock) activity were low.

In this study we derive an expression for the probability of a major earthquake following a possible foreshock near a major fault from the basic tenets of probability theory. This probability turns out to depend on the long-term probability of the mainshock, the rate of background seismicity along the fault, and some assumed characteristics of the relations between mainshocks and foreshocks. We then apply this expression to

Copyright 1991 by the American Geophysical Union.

Paper number 91JB00191.
0148-0227/91/91JB-00191\$05.00

the San Andreas fault system to develop short-term probabilities for possible earthquake warnings based on possible foreshocks.

2. MODELS FOR PROBABILITIES FROM FORESHOCKS

Because of the nature of seismicity along major fault systems such as the San Andreas fault, we have been led to address certain fundamental issues about the relationship between foreshocks and large earthquakes. These major faults illustrate in an extreme form the "maximum magnitude" model introduced by Wesnowsky *et al.* [1983], in which the frequency of the largest earthquakes on a fault zone is much higher than would be predicted by the extrapolation of the frequency-magnitude distribution for background earthquakes. For many parts of the San Andreas fault this is a straightforward consequence of the low level of present-day seismicity. For instance, along the Coachella Valley segment of the San Andreas fault (Figure 4) an extrapolation of present seismicity to higher magnitudes predicts a magnitude 7.5 earthquake every 2900 years, whereas the recurrence rate estimated from slip-rate data is 200–300 years.

This behavior implies that the large characteristic earthquakes on a fault zone are not simply the largest members of the total population of earthquakes there, but are somehow derived from a different population. Foreshocks to such events can thus reasonably be regarded as also being a separate class of events from the background earthquakes. A physical model that might underlie this is that some special failure process takes place before characteristic earthquakes, with an enhanced rate of small earthquakes and eventual failure on a large scale both being a result of it. It is of course also possible that no such process occurs; a moderate shock might, depending on the details of stress nearby, trigger only smaller events (in which case it is a mainshock) or larger ones (making it a foreshock), as suggested by Brune [1979]. There would then be no innate difference between background events and foreshocks; but we believe that it remains fruitful (as will be shown) to make at least a conceptual division.

That we make this division does not mean that there are any characteristics that can distinguish between foreshocks and other earthquakes; indeed, if there were, we would not have had to consider the second model above. We can only identify foreshocks, like aftershocks, by virtue of their association with a larger event; and, as our opening quotation suggests, for foreshocks such identification can only be retrospective. Such classification by association means that any particular shock might have been classified "incorrectly", and actually have been a background shock that just happened to fall close to a larger event. In our present state of knowledge this is unavoidable, and it may always remain so.

2.1. Zero-Dimensional Model

Starting from the assumption that foreshocks are a separate class of earthquake from background earthquakes, we can set out a formal probabilistic scheme for finding the probability of a large shock, given the occurrence of a possible foreshock. For clarity we begin with a "zero-dimensional" model, ignoring spatial variations, magnitude dependence, and other complications, which will be added in later sections. With these simplifications, a numerical example will illustrate the reasoning. Suppose that mainshocks occur every 500 years (on average), and that half of them have foreshocks (defined as being

within a day of the mainshock); then we expect a foreshock every 1000 years. If a comparable background earthquake occurs, on average, annually, we get 1000 background earthquakes per foreshock. If an earthquake occurs that could be either one, we then would assume the probability to be 10^{-3} that it is a foreshock, and so will be followed by a mainshock within a day. This is low, but still far above the background one-day probability of 5.5×10^{-6} .

For a formal treatment we begin by defining events (in the probability-theory meaning of the term):

B: A background earthquake has occurred.

F: A foreshock has occurred.

C: A large (characteristic) earthquake will occur.

As noted above, if a small background shock were to happen by coincidence just before the characteristic earthquake, we would certainly class it as a foreshock. Thus, *B* and *C* cannot occur together: they are disjoint. The same holds true for *B* and *F*: we can have a foreshock or a background earthquake, but not both.

The probability that we seek is the conditional one of *C*, given either *F* or *B*, because we do not know which has occurred. This is, by the definition of conditional probability,

$$P(C|F \cup B) = \frac{P(C \cap (F \cup B))}{P(F \cup B)} \quad (1)$$

Because *F* and *B* are disjoint, the probability of their union is the sum of the individual probabilities, allowing us to write the numerator of (1) as

$$P((C \cap F) \cup (C \cap B)) = P(C \cap F) + P(C \cap B) = P(C \cap F)$$

where the disjointness of *C* and *B* eliminates the $P(C \cap B)$ term. From the definition of conditional probability,

$$P(C \cap F) = P(F|C)P(C)$$

where $P(F|C)$ is the probability that a mainshock is preceded by a foreshock. Again using the disjointness of *F* and *B*, we can write the denominator as

$$P(F \cup B) = P(F) + P(B) \quad (2)$$

Because a foreshock cannot, by definition, occur without a mainshock, the intersection of *C* and *F* is *F*, and therefore

$$P(F) = P(F \cap C) = P(F|C)P(C) \quad (3)$$

We can use (2) and (3) to write (1) as

$$P(C|F \cup B) = \frac{P(F)}{P(F) + P(B)} = \frac{P(C)P(F|C)}{P(F|C)P(C) + P(B)} \quad (4)$$

For $P(B) \gg P(F|C)P(C)$ this expression is small (the candidate event is probably a background earthquake), while for $P(B) = 0$, the expression becomes equal to one: any candidate earthquake must be a foreshock.

The second form of expression in (4) is a function of three quantities, which in practice we obtain from very different sources. $P(B)$, the probability of a background earthquake, would be found from seismicity catalogs for the fault zone. $P(C)$, the probability of a characteristic earthquake, would be

found from calculations of the type presented by the *Working Group on California Earthquake Probabilities* [1988]. If we had a record of the seismicity before many such characteristic earthquakes, we could evaluate $P(F|C)$ (which we shall hereafter call Φ_{FC}) from it directly. (For this simple model, Φ_{FC} is just the fraction of large earthquakes preceded by foreshocks.) Of course, we do not have such a record, and so are forced to make a kind of reverse ergodic assumption, namely that the time average of Φ_{FC} over many earthquakes on one fault is equal to the spatial average over many faults. This may not be true, but it is for now the best we can do.

2.2. One-Dimensional Model

As a simple extension to the previous discussion, suppose that we have N "regions" and that C_i , B_i , and F_i denote the occurrence of an event in the i th region, with C (for example) now being the occurrence of a large earthquake in any possible region. These regions can be sections of the fault or (as we will see below) volumes in a multidimensional space of all relevant variables. The quantity of interest is now $P(C|F_i \cup B_i)$: we have a candidate foreshock in one region, and want the probability of a large earthquake starting anywhere. Assuming that the occurrences C_i are disjoint (the epicenter can only be in one place), we then have that the probability of a foreshock in the i th region can be written as

$$P(F_i) = \sum_{j=1}^N \Phi_{FC}(i, j) P(C_j) \quad (5)$$

where $\Phi_{FC}(i, j) = P(F_i|C_j)$. We may regard Φ_{FC} as the probability of a foreshock in region i given a large earthquake in region j . We call this the precurent probability because it refers to the probability of an event preceding a second one (not, it should be noted, with an implication of violated causality). As a simple example, we could take $\Phi_{FC}(i, j) = \alpha \delta_{ij}$, which would imply that large earthquakes are preceded by foreshocks only in the same region, and even then only a fraction α of them have foreshocks at all.

We can then easily revise (4) above to get the probability we seek; simply adding subscripts to the candidate event yields

$$P(C|F_i \cup B_i) = \frac{P(F_i) + P(C)P(B_i)}{P(B_i) + P(F_i)} \quad (6)$$

Equations (5) and (6) are the basic ones we shall use in the more general case. Equation (5) shows us how to compute the probability of a foreshock happening in the location of our candidate earthquake, by summing over all possible mainshocks. The use of the precurent probability Φ_{FC} is the key to this approach; we can (and in the next section shall) design it to embody our knowledge and assumptions about the relation between foreshocks and the earthquakes they precede. Having found the foreshock probability, we then use (6) to find the conditional probability of a large earthquake.

An important consequence of (5) is that we may sum over all possible foreshocks (again assuming disjointness) to get

$$P(F) = \sum_{i=1}^N \sum_{j=1}^N \Phi_{FC}(i, j) P(C_j) \quad (7)$$

giving us the overall probability of a foreshock somewhere in the total region. This must satisfy $P(F) = \alpha P(C)$, where α is

the fraction of mainshocks with foreshocks; this and equation (7) together constrain the normalization of Φ_{FC} .

Next to the probability level itself, the socially most interesting quantities would seem to be the chance of an alert being a false alarm, and the rate at which false alarms occur for a given probability level. The probability that an alert is a false alarm is $P(\bar{C}|F_i \cup B_i)$, which is just $1 - P(C|F_i \cup B_i)$; if we have a 10% chance of having a mainshock, we have a 90% chance of not having one. The rate of false alarms is equivalent to the probability of a false alarm happening in some given time, and this is just the probability that an alert is a false alarm times the probability of the event that triggers it, namely

$$\sum_{i=1}^N [1 - P(C|F_i \cup B_i)] [P(B_i) + P(F_i)]$$

As will be shown in section 4, we would in practice usually choose the probability of a mainshock given a small event, $P(C|F_i \cup B_i)$ to have a fixed value (e.g., 1%), which we denote by S , for all regions. This value of S then sets the value of $P(B_i)$ for the i th region; from (6), $P(B_i) = P(F_i)[(1 - S)/S]$, which makes the probability of a false alarm

$$\frac{(1 - S)}{S} \sum_{i=1}^N P(F_i) = \frac{(1 - S)}{S} \sum_{i=1}^N \sum_{j=1}^N \Phi_{FC}(i, j) P(C_j)$$

where we have used (5). For fixed S and Φ_{FC} this expression is proportional to $P(C_j)$ only: the rate of false alarms for a given probability depends only on the rate of mainshocks and not on the rate of background activity. In terms of the simple example at the beginning of section 2.1, fixing a probability level of 0.1% means that we would set the magnitude level of candidate events such that there would be 1000 background events for each actual foreshock; but the absolute rate of such background earthquakes (and thus of false alarms) is then determined only by the rate of foreshocks, and thus of mainshocks.

3. A MULTIDIMENSIONAL MODEL FOR FORESHOCKS

We now develop an expanded version of (5), which contains more variables. The first step is to define our events more thoroughly:

- B : A background earthquake has occurred at coordinates $(x_0 \pm e_0, y_0 \pm e_0)$, during the time period $[t, t + \delta_0]$, with magnitude $M \pm \mu$. (All of the quantities e_0 , δ_0 , and μ are small and are included because we will be dealing with probability density functions; as will be seen below, they cancel from the final expression).
- F : A foreshock has occurred, with the same parameters as in event B .
- C : A major earthquake will occur somewhere in the region of concern, which we denote by A_C (also using this variable for the area of this region). This earthquake will happen during the time period $[t + \Delta, t + \Delta + \delta_1]$, with magnitude between M_C and $M_C + \mu_C$.

We assume that we are computing the probability at some time in the interval $(t + \delta_0, t + \Delta)$; the possible foreshock has happened, but the predicted mainshock is yet to come.

3.1. Rate Densities of Earthquake Occurrence

We begin by defining a rate of occurrence for the background seismicity (in the literature on point processes this would be

called an intensity, a term we avoid because of existing seismological usage). This rate (or, strictly speaking, rate density) we call $\Lambda(x, y, M)$; it is such that the probability of B is

$$P(B) = \delta_0 \int_{x_0 - \epsilon_0}^{x_0 + \epsilon_0} dx \int_{y_0 - \epsilon_0}^{y_0 + \epsilon_0} dy \int_{M - \mu}^{M + \mu} dM \Lambda(x, y, M) \quad (8)$$

By not making Λ dependent on the time t we make the occurrence of background earthquakes into a Poisson process. If we assume that at any location the Gutenberg-Richter frequency-magnitude relation holds, we may write

$$\Lambda(x, y, M) \equiv \Lambda_s(x, y) e^{-\beta(x, y)M} \quad (9)$$

where β is 2.3 times the usual b value. (While common rather than natural logarithms are conventional in this area, they lead to messier expressions, and we have therefore not used them). If β is constant over a region of area A , and during a time interval T the cumulative number of earthquakes of magnitude M or greater is given by the usual formula

$$N(M) = 10^{a-bM} \quad (10)$$

then, since the expected value of $N(M)$ is

$$E[N(M)] = T \int_M^\infty dM \iint_A dx dy \Lambda_s(x, y) e^{-\beta M} \quad (11)$$

we have that $\Lambda_s = (10^a \beta) / (AT)$ for Λ_s constant within the region.

Similarly, we can define a rate density for the occurrence of large earthquakes,

$$\Omega(x, y, M, t) \equiv \Omega_s(x, y, t) e^{-\beta'(x, y)M} \quad (12)$$

where β' is 2.3 times the b value for these events. In this case, we introduce a dependence on time t because the occurrence of large earthquakes is often formulated as a renewal process [e.g., Nishenko and Buland, 1987], with time being measured relative to the last earthquake. The probability of C is then

$$P(C) = \delta_1 \iint_{A_C} dx dy \int_{M_C}^{M_C + \mu_C} dM \Omega(x, y, M, t + \Delta) \quad (13)$$

where A_C is the area of concern, i.e., the particular segment of a fault.

For lack of better information we would usually take Ω_s to be a constant, but we could choose to make it spatially varying. Such variation could include increases near fault jogs and terminations if we think that rupture nucleation is more likely there, or a proportionality to Λ_s if we suspect that background earthquakes are (on the average) the likely triggers of large ones (both issues are discussed in section 3.2.2). For Ω_s constant, we have that

$$\Omega_s = \frac{P(C)\beta'}{A_C \delta_1 e^{-\beta' M_C} (1 - e^{-\beta' \mu_C})} \quad (14)$$

Note that while we have regarded both A and A_C as two-dimensional regions (and hence also as the areas of such

regions), we may in fact make them three-dimensional or one-dimensional if we so choose, making sure that we adjust the numbers of the integrals in (8) and (13) accordingly. The one-dimensional model is easiest to develop analytical expressions for, and may be an adequate approximation for the case of a long fault zone. In this case, of course, we need to project the background seismicity (out to some distance away) onto the fault zone.

3.2. Computation of the Foreshock Probability

We are now in a position to write the formal expression for the foreshock probability $P(F)$ in the same way as was done in (5) for the discrete one-dimensional case. In this case, Φ_{FC} becomes a density function over all the variables involved, its value indicating with what relative frequency foreshocks with different parameters occur before mainshocks with particular ones. Instead of a single sum, as in (5), we have a multiple integral:

$$P(F) = \int_t^{t+\delta_0} dt \int_{x_0 - \epsilon_0}^{x_0 + \epsilon_0} dx \int_{y_0 - \epsilon_0}^{y_0 + \epsilon_0} dy \int_{M - \mu}^{M + \mu} dM \int_{t+\Delta}^{t+\Delta+\delta_1} dt' \iint_{A_C} dx' dy' \int_{M_C}^{M_C + \mu_C} dM' \Phi_{FC}(t, t', x, y, x', y', M, M') \Omega_s(x', y', t') e^{-\beta'(x', y')M'} \quad (15)$$

Of these eight integrals, the last four are the integration of the precurent probability density times the density of mainshock occurrence over the space of possible mainshocks and are the equivalent of the sum in (5). But this gives only the rate density for foreshocks, which must in turn be integrated over the space of the candidate event (the first four integrals) to produce the actual probability $P(F)$.

Equation (15) is clearly quite intractable as it stands. To render it less so, we assume that we can separate the behaviors of $P(F)$ in time, magnitude, and location. This implies the following assumptions:

- 1: β' does not depend on x' or y' .
- 2: Over the range of integration, Ω_s does not depend on t' .
- 3: The functional forms of the precurent probability density for time, space, and magnitude are independent, so that we can write the function as the product of the marginal distributions:

$$\Phi_{FC} = \Phi_s(x, y, x', y') \Phi_t(t, t') \Phi_m(M, M')$$

Of these assumptions, the third seems the least likely to be valid, since the dependence on both distance and time might be correlated with the magnitude of either the mainshock or the candidate foreshock. The most likely correlation, with mainshock magnitude, does not matter very much, since our range of integration of this variable is small.

These assumptions made, we can divide the integral in (15) into a product of three integrals (in space, time, and magnitude):

$$P(F) = \int_t^{t+\delta_0} dt \int_{t+\Delta}^{t+\Delta+\delta_1} dt' \Phi_t(t, t') \int_{M - \mu}^{M + \mu} dM \int_{M_C}^{M_C + \mu_C} dM' \Phi_m(M, M') e^{-\beta' M'} \int_{x_0 - \epsilon_0}^{x_0 + \epsilon_0} dx \int_{y_0 - \epsilon_0}^{y_0 + \epsilon_0} dy \iint_{A_C} dx' dy' \Phi_s(x, y, x', y') \Omega_s(x', y') \quad (16)$$

3.3. Functional Forms for the Foreshock Density

To evaluate the integrals in (16), we need to know the three precurrent probability densities Φ_t , Φ_s , and Φ_m . Our expressions for these incorporate our knowledge and assumptions about foreshocks. In the following sections, we describe in some detail what is known about the temporal, spatial, and magnitude dependences of foreshocks. From these data, we find functions for the relevant Φ ; these functions must include both the actual dependence on the variables and a normalization. The nature of the normalization can be seen if we imagine extending the range of the first four integrals in (15) to cover all possible foreshocks (however we chose to define them); the resulting $P(F)$ must then be equal to $\alpha P(C)$, where α is, as for the one-dimensional model, the fraction of mainshocks preceded by foreshocks. In deriving our expressions we have aimed for simplicity rather than attempting to find a function that can be shown to be statistically optimal.

3.3.1 Time. Most foreshocks occur just before the mainshock. An increase in earthquake occurrence above the background rate has only been seen for a few days [Jones, 1984; 1985; Reasenberg, 1985] to a week [Jones and Molnar, 1979] before mainshocks. For 26% of Californian mainshocks, the foreshocks are most likely to occur within 1 hour of the mainshock; the rate of foreshock occurrence before mainshocks (Figure 1) varies with the t^{-1} type behavior also seen in Omori's law for aftershocks [Jones, 1985; Jones and Molnar, 1979]. This variation can be well fit by the function that Reasenberg and Jones [1989] found for California aftershock sequences:

$$\Phi_t(t, t') = \frac{N_t}{t' - t + c} \quad (17)$$

where t is the foreshock time and t' the mainshock time; c is a constant, found by Reasenberg and Jones [1989] to be 200 s for aftershocks. The relevant integral from (16) is then

$$\int_{t-\delta_0}^{t+\delta_0} dt \int_{t-\Delta}^{t+\Delta} dt' \Phi_t(t, t') = \delta_0 N_t \ln[1 + \delta_1/(\Delta + c)] \quad (18)$$

where we have assumed that δ_0 (the uncertainty of the time of

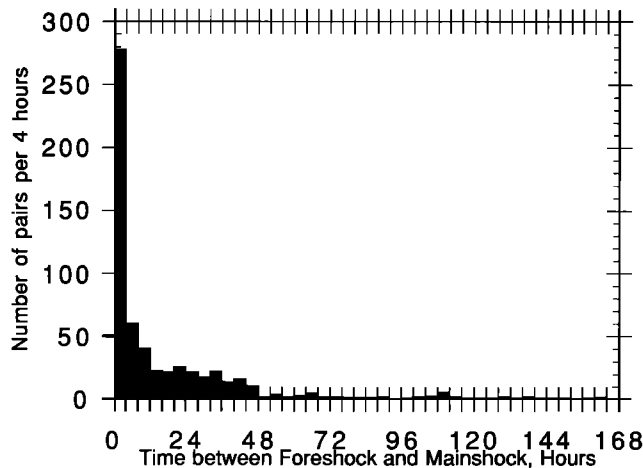


Fig. 1. The number of foreshock-mainshock pairs recorded in southern California versus the time between foreshock and mainshock in hours for foreshocks $M \geq 2.0$ and mainshocks $M \geq 3.0$ recorded between 1932 and 1987.

the candidate earthquake) is small. The normalization is determined by the requirement that

$$\int_{t-\delta_0}^{t+\delta_0+t_W} dt' \Phi_t(t, t') = 1 \quad (19)$$

where t_W is the total time window within which we admit preceding earthquakes to be foreshocks. This then gives

$$\begin{aligned} \int_{t-\delta_0}^{t+\delta_0} dt \int_{t-\Delta}^{t+\Delta} dt' \Phi_t(t, t') &= \delta_0 \frac{\ln[1 + \delta_1/(\Delta + c)]}{\ln[1 + t_W/(\delta_0 + c)]} \\ &\equiv \delta_0 I_t(\Delta, \delta_1) \end{aligned} \quad (20)$$

where, with an eye to future simplifications, we have separated out the δ_0 term. Note that (17) predicts a finite rate for all times, whereas the assumption of a limited time window automatically forces the rate to fall to zero beyond some time; we can easily modify Φ_t to allow for this.

3.3.2 Location. Foreshocks not only occur close in time to the mainshock, but are also nearby in space. Jones and Molnar [1979] found that epicenters of mainshocks ($M \geq 7$) and their foreshocks in the National Earthquake Information Center (NEIC) catalog were almost all within 30 km of each other, approximately the location error for the NEIC catalog. Jones [1985], with the more accurate locations of the California Institute of Technology (Caltech) catalog, found that epicenters of mainshocks ($M \geq 3$) and their foreshocks were almost all within 10 km of each other; this result also held for foreshocks of $M \geq 5$ mainshocks within the San Andreas system [Jones, 1984] if relative relocations were used. Even the largest foreshocks ($M \geq 6$ at Mammoth Lakes and Superstition Hills) have had epicenters within 10 km of the epicenters of their mainshocks.

We have assembled a data set of sequences with high-quality locations to examine the dependence of the distance between foreshocks and mainshocks on the magnitudes of the earthquakes. This data set includes all foreshock-mainshock pairs with $M_{fore} \geq 2.5$ and $M_{main} \geq 3.0$ recorded in southern California since 1977 (the start of digital seismic recording), and several sequences relocated in special studies, with relative location accuracy of at least 1 km. Figure 2 shows the distance between foreshock and mainshock versus magnitude of the mainshock (2a) and magnitude of the foreshock (2b). The epicentral separation between foreshock and mainshock does not correlate strongly with either magnitude. Rather, the data seem to group into two classes: foreshocks that are essentially at the same site as their mainshock (<3 km) and foreshocks that are clearly separated from their mainshocks. Only foreshocks to larger mainshocks ($M_{main} \geq 5.0$) occur at greater epicentral distances (5–10 km). Of these spatially separate foreshocks some (but not all) ruptured towards the epicenter of the mainshock (the rupture zones are shown by the ovals in Figure 2). The greatest reported distance between foreshock and mainshock epicenters is 8.5 km; the greatest reported distance between foreshock rupture zone and mainshock epicenter is 6.5 km. It would therefore seem that, whatever other behavior Φ_s may have, it can be taken to be zero for distances greater than 10 km.

It is possible (and allowed for in our choice of variables for Φ_s) for foreshocks to be preferentially located in some sections

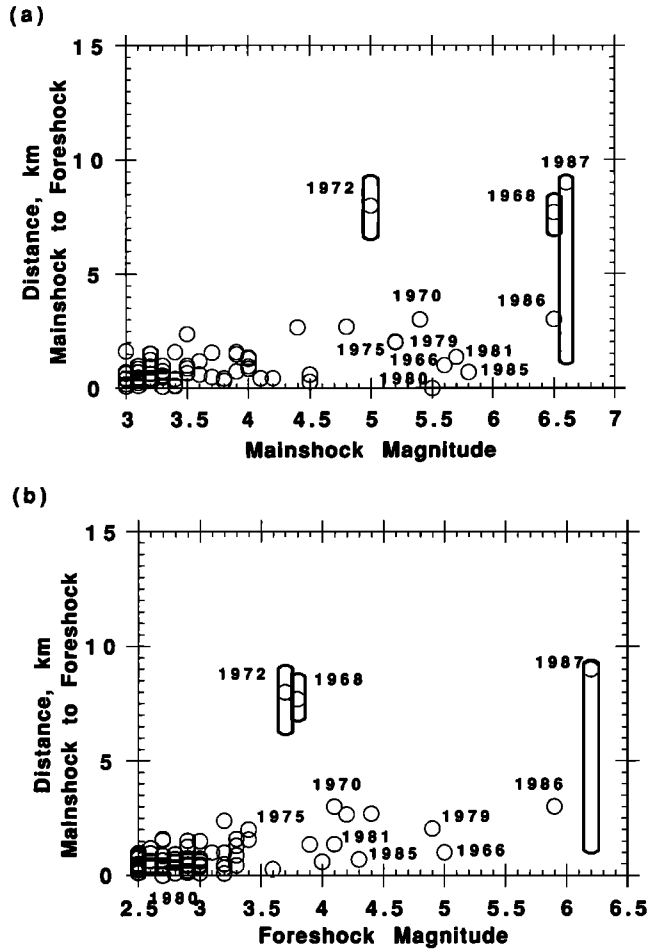


Fig. 2. Distance between foreshock and mainshock epicenters versus the (a) magnitude of the mainshock and (b) magnitude of the foreshock for foreshock-mainshock sequences (foreshocks $M \geq 2.5$ and mainshocks $M \geq 3.0$) recorded in the Caltech catalog between 1977 and 1987. Sequences that have been relocated in special studies are also plotted and include 1966 Parkfield, 1968 Borrego Mountain, 1970 Lytle Creek, 1972 Bear Valley, 1975 Haicheng ($M = 7.3$), 1975 Galway Lakes ($M = 5.2$), 1979 Homestead, 1980 Livermore, 1981 Westmoreland, 1985 Kettleman Hills, 1986 Chalfant Valley, and 1987 Superstition Hills. For the three foreshock sequences with known rupture zones the distance range of foreshock rupture zone to the mainshock epicenter is shown by the elongated ovals; the circles inside these show the distance for the foreshock epicenter.

of major faults. Jones [1984] suggested that foreshocks are more common at areas of complication along faults; this would require that either $P(C)$ or Φ_s (or both) be larger at such places. An increase of $P(C)$ would be in accordance with the notion that epicenters of mainshocks are mostly at such points [King and Nabelek, 1985; Bakun *et al.*, 1986]. While this seems like a valid refinement, in practice differentiating between the many possible complex sites and the "smooth" parts of the fault would require gridding at the kilometer scale, a level of detail that does not seem justified by our present level of knowledge. One further choice would be to make Φ_s proportional to the local rate of background activity Λ_s , thus asserting that most mainshocks with foreshocks occur in areas with high background seismicity. The data on foreshocks to moderate earthquakes in California [Jones, 1984] does not support this: while the fraction of earthquakes with foreshocks does vary by region, it does not appear to be related to background seismicity.

city. For example, the Calaveras fault in central California has a relatively high rate of background activity and no foreshocks.

Foreshocks and mainshocks thus clearly occur close together in space, within 10 km of each other in all resolvable cases - but show no other clear dependence on location. We therefore have made Φ_s depend only on ρ , the distance between candidate foreshock and possible mainshock epicenters ($\rho = [(x - x')^2 + (y - y')^2]^{1/2}$). The condition for Φ_s to be properly normalized is

$$\iint_{A_C} dx dy \iint_{A_C} dx' dy' \Phi_s(x, y, x', y') \Omega_s(x', y') = \iint_{A_C} dx' dy' \Omega_s(x', y') \quad (21)$$

which in general can be done only numerically, even for Ω_s constant and Φ_s having a simple dependence on ρ . If, however, we make the simplification, mentioned in section 3.1, of making our spatial integrals one-dimensional (with A_C then being the length of the fault), assume Ω_s constant, and make Φ_s constant for $\rho \leq \rho_w$ and zero for larger ρ , we find that Φ_s is

$$\frac{1}{\rho_w(1 - \rho_w^2/4A_C)} \quad \text{if } \rho \leq \rho_w \\ 0 \quad \text{if } \rho > \rho_w \quad (22)$$

We use $\rho_w = 10$ km to agree with the data presented above. Then, provided that the location x_0 of the candidate earthquake is more than a distance ρ_w from an end of the fault zone and that $\Omega_s(x')$ is constant over a distance $2\rho_w$, the integral needed in (16) is

$$\int_{x_0 - \rho_w}^{x_0 + \rho_w} dx \int_{A_C} dx' \Phi_s(x, x') \Omega_s(x) = 2e_0 \frac{\Omega_s(x_0)}{1 - \rho_w^2/4A_C} \equiv 2e_0 I_s(x_0) \quad (23)$$

where we have defined I_s in a parallel way to I_f ; the dependence on x_0 comes through the dependence on the value of Ω_s near the candidate earthquake.

3.3.3. Magnitude. The functional form for $\Phi_m(M, M')$ is probably the least certain part of Φ_{FC} . Plots of the difference in foreshock and mainshock magnitudes with a uniform magnitude threshold for foreshocks and mainshocks [e.g., Jones, 1985] show the magnitude difference to be a negative exponential distribution. However, to consider all possible foreshocks to a given mainshock, the completeness threshold for the foreshocks should be much lower than for the mainshocks. A bivariate plot of foreshock and mainshock magnitudes for all recorded foreshocks in southern California (Figure 3) suggests that for any given narrow range of mainshock magnitude, foreshock magnitudes close to that of the mainshock are more common; however, for the larger mainshock magnitudes of interest here, the (admittedly sparse) data suggest that all foreshock magnitudes are equally likely for given mainshock magnitude.

Because of the simplicity of this last assumption, we have used it here by making Φ_m constant; we set $\Phi_m(M, M') \equiv N_m$, a normalizing factor. The normalization of Φ_m is in general set by

$$\int_{M_B}^{M'} \int_{M_D}^{M'} \Phi_m(M, M') dM dM' = \alpha \int_{M_B}^{\infty} e^{-\beta M'} dM' \quad (24)$$

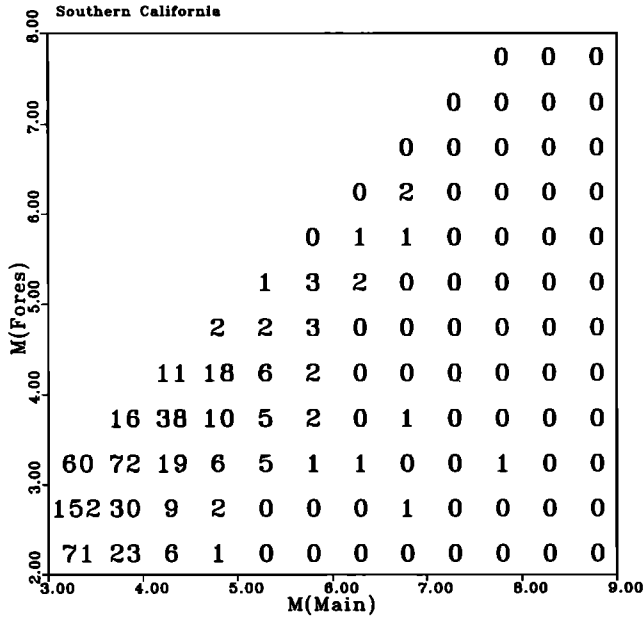


Fig. 3. The number of foreshock-mainshock pairs in half unit of magnitude bins for the magnitudes of foreshock and mainshock. Data included all $M \geq 2.0$ foreshocks and $M \geq 3.0$ mainshocks recorded between 1932 and 1987 in southern California.

Equation (24) says that if we look before all mainshocks with magnitudes greater than M_B for foreshocks above a cutoff magnitude of M_D , we find that a fraction α of the mainshocks have foreshocks. Note that we have chosen to normalize Φ_f and Φ_s to integrate to 1, so Φ_m contains the information about the total fraction of mainshocks with foreshocks.

Making Φ_m constant implies that the fraction of mainshocks preceded by foreshocks will increase as the magnitude threshold for foreshocks decreases. This is consistent with reported foreshock activity, since the data suggest that foreshocks are relatively common before major strike-slip earthquakes. Jones and Molnar [1979] found that 30% of the $M \geq 7.0$ earthquakes occurring outside of subduction zones were preceded by foreshocks in the NEIC catalogue ($M \geq 4.5-5.0$) and almost 50% had foreshocks $M \geq 2$ reported in the literature. Jones [1984] showed that half of the $M \geq 5.0$ strike-slip earthquakes in California were preceded by $M \geq 2.0$ foreshocks. (Foreshocks were less common on thrust faults.)

For Φ_m constant and equal to N_m , (24) implies that

$$N_m = \frac{\alpha\beta'}{1 + \beta'(M_B - M_D)} \quad (25)$$

The data presented by Jones [1984], with $M_B = 5.0$ and $M_D = 2.0$, gave α equal to 0.5 for strike-slip earthquakes. Adopting this value, with a β' of 2.3, gives $N_m = 0.15$. A consequence of taking Φ_m constant is then that all earthquakes should have foreshocks within 6.5 units of magnitude of the mainshock. Holding Φ_m constant for all M would of course lead to the absurd result that more than 100% of mainshocks have foreshocks within, say, 8 magnitude units. For the smaller range of magnitudes considered here a constant Φ_m does not present any difficulties.

The integral needed for (16) is then

$$\int_{M-\mu}^{M+\mu} dM \int_{M_C}^{M_C+\mu_C} dM' \Phi_m(M, M') e^{-\beta'M'} = 2\mu N_m e^{-\beta'M_C} \frac{1 - e^{-\beta'\mu_C}}{\beta'} \equiv 2\mu I_m(M_C, \mu_C) \quad (26)$$

where we have assumed μ small, and again separated it out from the rest of the expression.

3.4. Mainshock Probability

We now can combine the integrals in (18), (23), and (26) into (16) to get the foreshock probability:

$$P(F) = 4\delta_0\mu e_0 I_f I_s I_m$$

Solving the integral in (8) for the background event gives

$$P(B) = 4\delta_0\mu e_0 \Lambda_s(x_0) e^{-\beta M}$$

We substitute these values of the background and foreshock probabilities into (6) to obtain:

$$P(C|F \cup B) = \frac{I_s I_f I_m}{I_s I_f I_m + \Lambda_s(x_0) e^{-\beta M}} \quad (27)$$

The candidate earthquake errors δ_0 , e_0 , and μ have canceled out.

For making calculations, it is also useful to set I_f equal to 1 (solve for the probability in a fixed time interval) and (for the case of a linear fault) take I_s in (23) to be equal to $\Omega_s(x_0)$. If we take Ω_s to be constant and combine (14) and (26), we find that the dependence on M_C and μ_C cancels out, and we are left with

$$P(C|F \cup B) = \frac{(N_m P(C)/A_C \delta_1)}{(N_m P(C)/A_C \delta_1) + \Lambda_s(x_0) e^{-\beta M}} \quad (28)$$

4. APPLICATION TO THE SAN ANDREAS FAULT SYSTEM, CALIFORNIA

We now have an expression for the conditional probability of a characteristic earthquake on a fault segment given the occurrence of an earthquake that is either a background event or a foreshock. To evaluate this, we need the long-term probability of the characteristic mainshock (the terms involving the actual magnitude of the characteristic earthquake have canceled out), the length of the fault segment, and the rate density of background seismicity for that segment. To show how this works, we now apply this to the San Andreas fault system in California, because the long-term probabilities for characteristic earthquakes that we need have been estimated for the major faults of this system, the San Andreas, Hayward, San Jacinto and Imperial faults. This was first done by Lindh [1983] and Sykes and Nishenko [1984], and more recently by the Working Group on California Earthquake Probabilities [1988], hereafter referred to as WGCEP-88.

Our division of the fault into segments and our values of $P(C)$ for each segment come largely from WGCEP-88. One exception is that the lengths of the Southern Santa Cruz Mountains and the San Francisco Peninsula segments have been altered to match the rupture zone of the 1989 Loma Prieta earth-

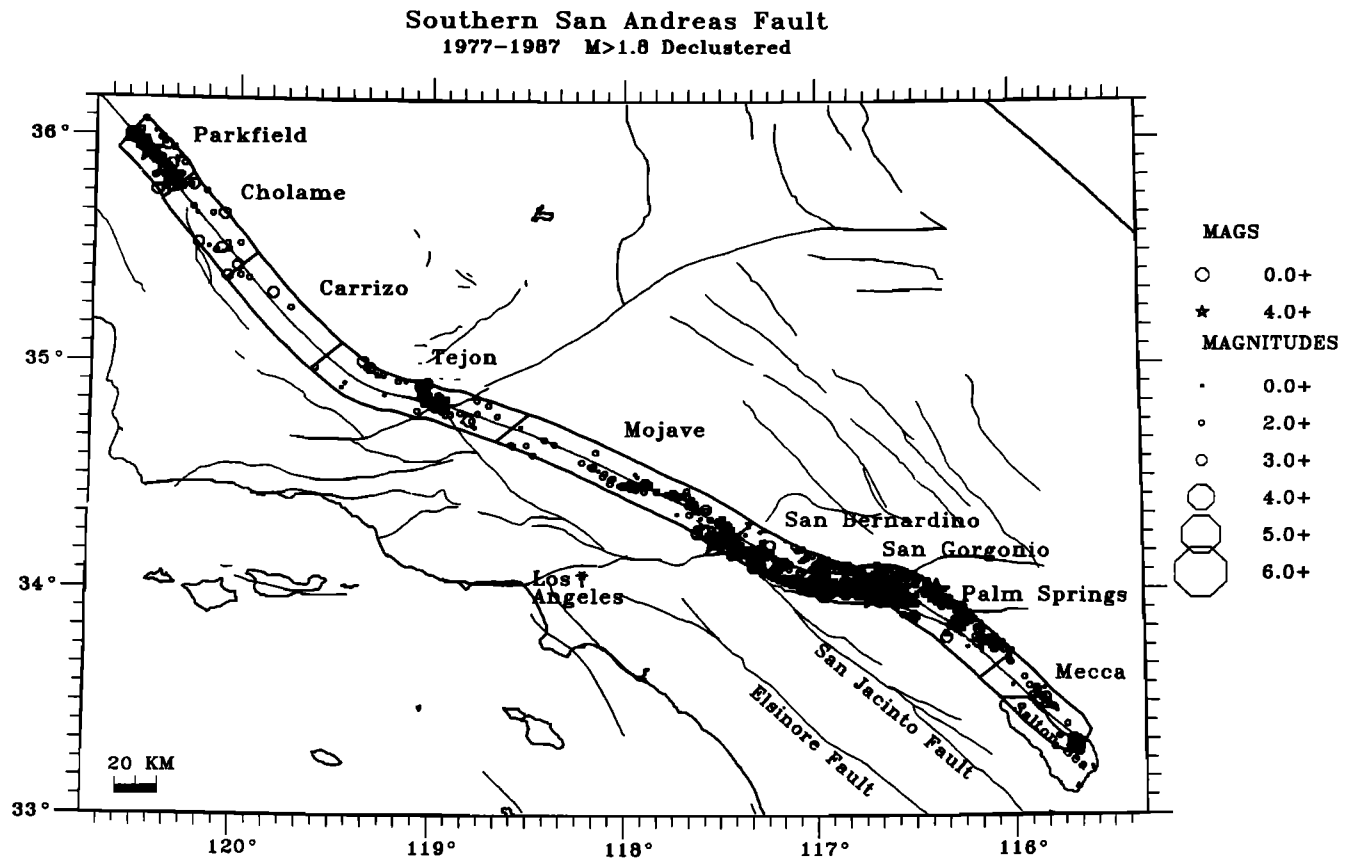


Fig. 4. A map of $M \geq 1.8$ declustered earthquakes located within 10 km of the southern San Andreas fault recorded in the Caltech catalog between 1977 and 1987 and (for Parkfield) the CALNET catalog between 1975 and 1989.

quake (A. Lindh, personal communication, 1990). We took $P(C)$ to be constant along each segment; as noted in section 3.2.1, we have not tried to include the possibility that nucleation points (and higher values of $P(C)$) are more likely at points of complication. We have also not altered the distribution of $P(C)$ to account for any possible relationship between nucleation point and level of background activity.

The rate density for the background seismicity is determined from the microearthquakes recorded between 1977 and 1987 by the Caltech/U.S. Geological Survey Southern California Seismic Network [Given *et al.*, 1988] for southern California and between 1975 and 1989 by CALNET, the U.S. Geological Survey Central California Seismic Network (P. Reasenberg, personal communication, 1990), for northern California. Background seismicity can be defined in many ways; it is important in this application that it be defined in the same way as the foreshocks will be. Because foreshocks can be up to 10 km from their mainshock (Figure 2), background seismicity up to 10 km from the surface trace of the San Andreas fault is included in the background rate.

Another issue is how to handle temporal clustering in the catalog. We assume that if an earthquake of $M = 6$ (for instance) were to occur on the southern San Andreas fault with an aftershock sequence, we will only evaluate the probability that the $M = 6$ earthquake is a foreshock, and not individually determine the probabilities that the $M = 6$ and each of its aftershocks is a foreshock and then sum them. For consistency we therefore want to determine the background seismicity using a

catalog from which aftershock sequences and swarms have been removed. In such a declustered catalogue, sequences are recognized by some algorithm and replaced in the catalogue with one event at the time of the largest earthquake in the sequence, which is given a magnitude equivalent to the summed moment of all the earthquakes in the sequence. To produce our declustered catalogs, we used the algorithm of Reasenberg [1985].

The resulting background seismicity within 10 km of the faults is shown in Figures 4-6. It is clear from these that the rate of background seismicity can vary significantly within the fault segments defined by WGCEP-88. For example, the Coachella Valley segment of the San Andreas includes the active region around Desert Hot Springs (including a $M = 6.5$ event in 1948) and a very quiet region (near the Salton Sea) where the largest earthquake in 55 years has been $M = 3.5$. To account for this variation, we have divided some of the WGCEP-88 segments into smaller regions, which are shown in Figures 4-6 and listed in Table 1.

Table 1 provides the data needed for each segment. To use (28) we also need the time period δ_1 , which we set to 3 days (1.09×10^5 s), to match the recent usage of the U.S. Geological Survey and the California's Governor's Office of Emergency Services in issuing earthquake advisories. Alert levels for such advisories are defined to correspond to certain probabilities; the magnitudes of earthquakes needed to trigger those alert levels can then be computed from (28), and are also given in Table 1. Figure 7 shows the probability as a function of the magnitude of

Northern San Andreas Fault System Declustered 1975-1989

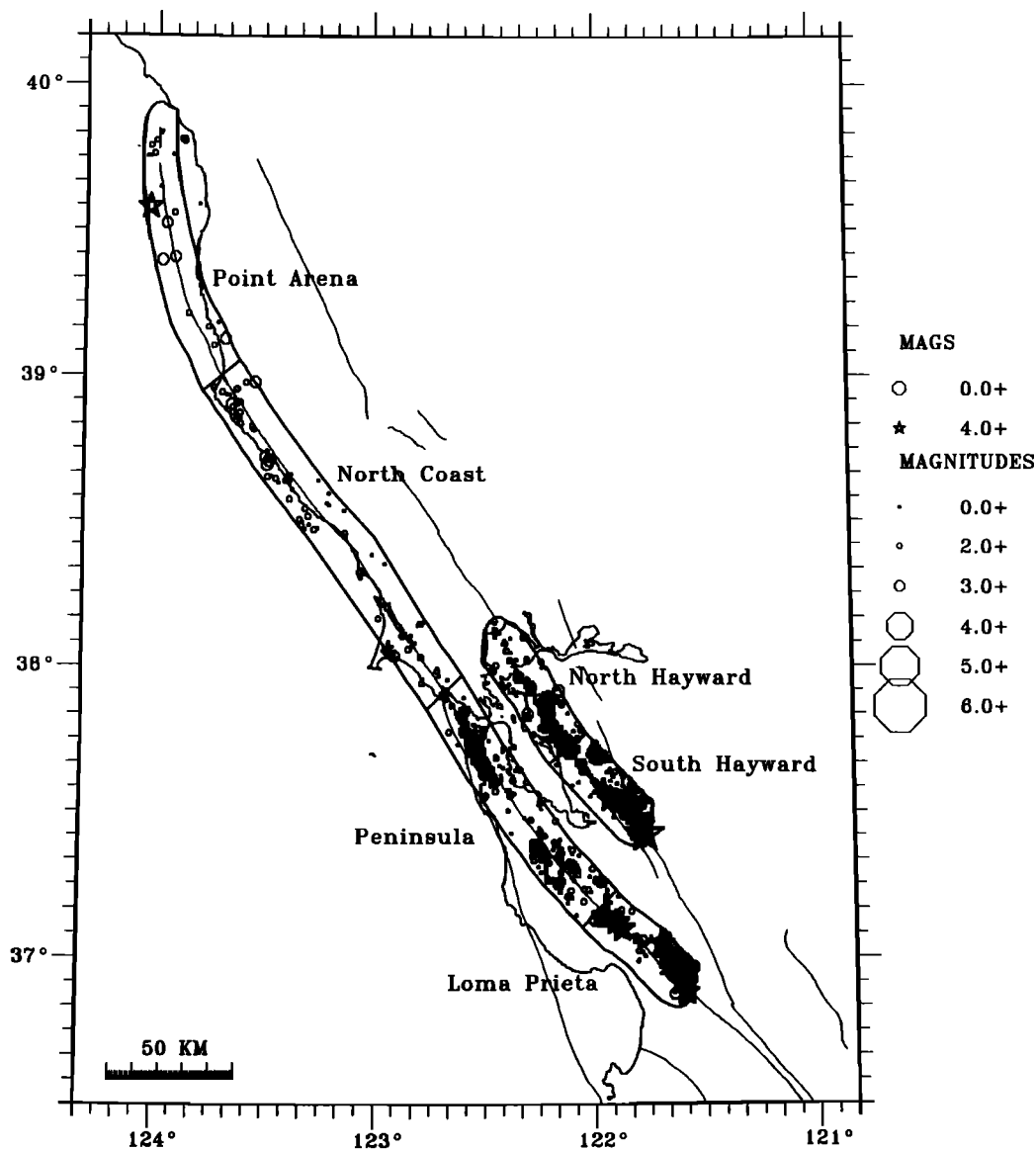


Fig. 5. A map of $M \geq 1.8$ declustered earthquakes located within 10 km of the San Jacinto fault recorded in the Caltech catalog between 1977 and 1987.

the candidate earthquake for each segment.

We have treated the Parkfield segment in two different ways. In the Table 1 listing for Parkfield, we treat it in the same way as the other segments, regarding the foreshock as equally likely anywhere along the segment, and taking $P(C)$ from WGCEP-88. These assumptions give short-term probabilities much lower than those estimated by *Bakun et al.* [1987] for the Parkfield earthquake prediction experiment. *Bakun et al.* [1987] used a somewhat different methodology and also used different assumptions in two areas: their value of $P(C)$ is 1.5 times that of WGCEP-88, and they assume that the foreshock will be located in a small region under Middle Mountain, making a smaller area for defining background seismicity. (Their assumption that 50% of Parkfield mainshocks will be preceded by foreshocks agrees with our choice in section 3.2.3). For a better comparison we have used the *Bakun et al.* assumptions to deter-

mine short-term probabilities with our methodology and given these in Table 1 as Middle Mountain probabilities. These remain lower than the *Bakun et al.* results; for example, a magnitude 1.5 shock gives a probability of 0.1% from our methodology and 0.68% (Level D alert) according to *Bakun et al.*

As with the long-term probabilities of major earthquakes, these short-term foreshock-based probabilities are better seen as a means of ranking the relative hazard from different sections of the faults than as highly accurate absolute estimates. The probabilities are as uncertain as the data used to calculate them, which in some cases are uncertain indeed. For example, the values of $P(C)$ found by WGCEP-88 are up to a factor of 4 larger than those found by *Davis et al.* [1989]; this would lead to similarly large differences in the short-term probabilities.

The relative short-term probabilities for different segments shown in Table 1 and Figure 7 are strongly affected by both the

San Jacinto and Imperial Faults
1977-1988 $M > 1.8$ Declustered

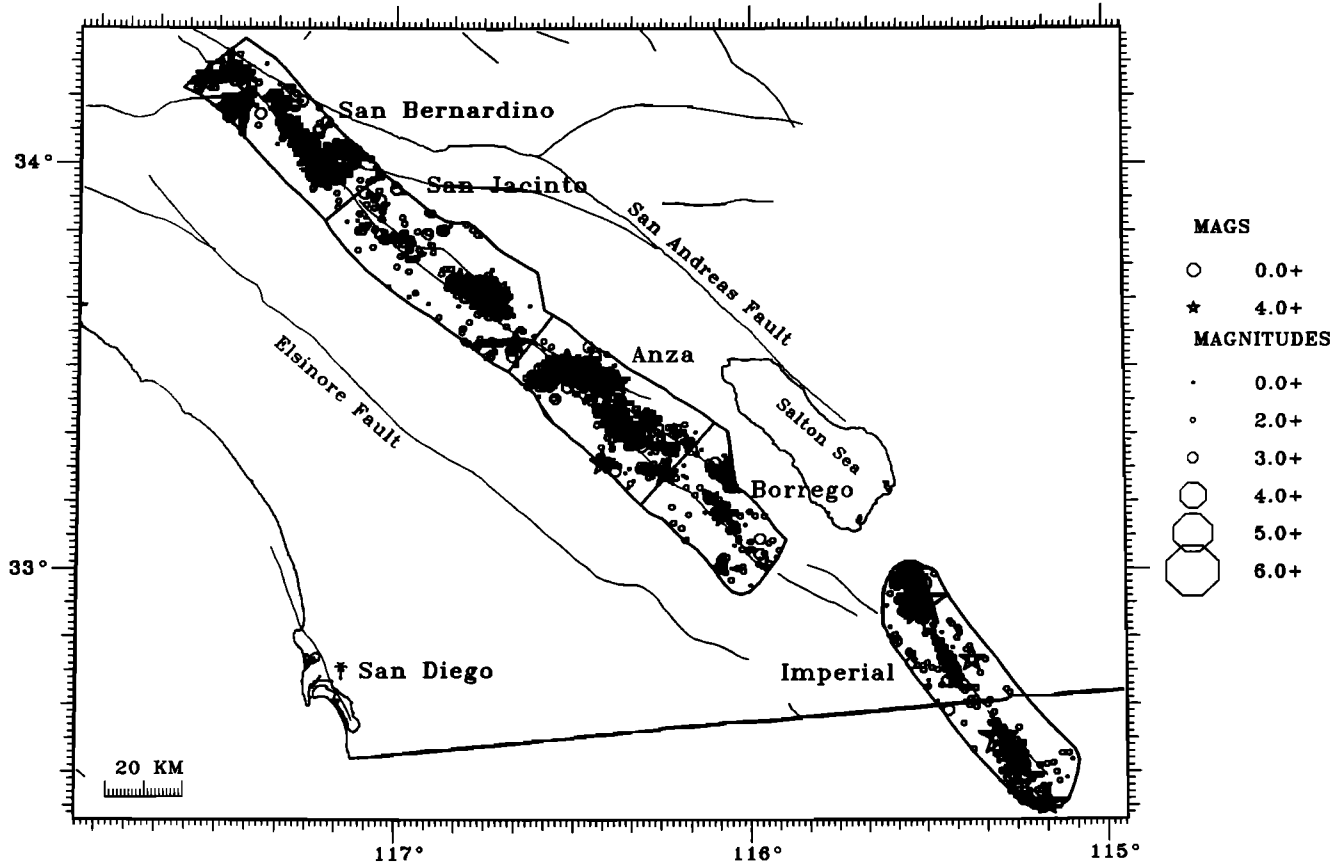


Fig. 6. A map of $M \geq 1.5$ declustered earthquakes located within 10 km of the northern San Andreas and Hayward faults recorded in the CALNET catalog between 1975 and 1989.

TABLE 1. Parameters and Magnitude Levels

Segment	Length, km	P_C /3 days	a	b	Λ_s , events/km s	β	Magnitude for		
							0.1%	1%	10%
<i>San Andreas Fault</i>									
Mecca	60	1.1×10^{-4}	3.67	0.95	4.91×10^{-7}	2.18	3.1	4.2	5.3
Palm Springs	50	1.1×10^{-4}	4.00	0.97	1.29×10^{-6}	2.23	3.5	4.5	5.6
San Gorgonio	60	5.5×10^{-5}	4.46	0.94	2.99×10^{-6}	2.16	4.2	5.3	6.4
San Bernardino	40	5.5×10^{-5}	3.95	0.92	1.36×10^{-6}	2.12	4.0	5.1	6.2
Mojave	100	8.2×10^{-5}	3.85	0.90	4.22×10^{-7}	2.07	3.3	4.4	5.6
Tejon	100	2.7×10^{-5}	3.49	0.88	1.80×10^{-7}	2.02	3.7	4.9	6.1
Carrizo	60	2.7×10^{-5}	2.58	1.03	4.32×10^{-8}	2.37	2.6	3.6	4.6
Cholame	50	8.2×10^{-5}	2.87	0.83	8.15×10^{-8}	1.91	2.3	3.6	4.8
Parkfield	35	8.2×10^{-4}	4.17	0.87	1.79×10^{-6}	2.00	2.5	3.6	4.8
Middle Mountain	20	1.2×10^{-3}	3.40	0.74	4.52×10^{-7}	1.70	1.5	2.9	4.3
Loma Prieta	50	8.2×10^{-5}	4.41	1.01	2.52×10^{-6}	2.32	3.4	4.4	5.4
Peninsula	100	5.5×10^{-5}	4.57	1.15	2.08×10^{-6}	2.64	3.3	4.2	5.1
North Coast	150	1.4×10^{-5}	3.26	0.88	6.09×10^{-8}	2.02	3.7	4.9	6.1
Point Arena	100	1.4×10^{-5}	2.95	0.69	1.40×10^{-8}	1.59	3.8	5.3	6.8
<i>San Jacinto Fault</i>									
San Bernardino	50	5.5×10^{-5}	4.58	0.98	4.94×10^{-6}	2.25	4.0	5.0	6.1
San Jacinto	65	2.7×10^{-5}	4.49	1.01	3.18×10^{-6}	2.32	4.1	5.1	6.1
Anza	50	8.2×10^{-5}	4.57	0.95	4.68×10^{-6}	2.18	3.9	5.0	6.1
Borrego	40	1.4×10^{-5}	4.05	0.99	1.84×10^{-6}	2.28	4.0	5.0	6.1
<i>Hayward Fault</i>									
North Hayward	60	5.5×10^{-5}	4.24	0.99	1.39×10^{-6}	2.28	3.5	4.5	5.5
South Hayward	50	5.5×10^{-5}	4.41	1.01	2.52×10^{-6}	2.32	3.6	4.6	5.6
<i>Imperial Fault</i>									
Imperial	50	1.4×10^{-4}	4.59	0.96	4.95×10^{-6}	2.21	3.6	4.7	5.8

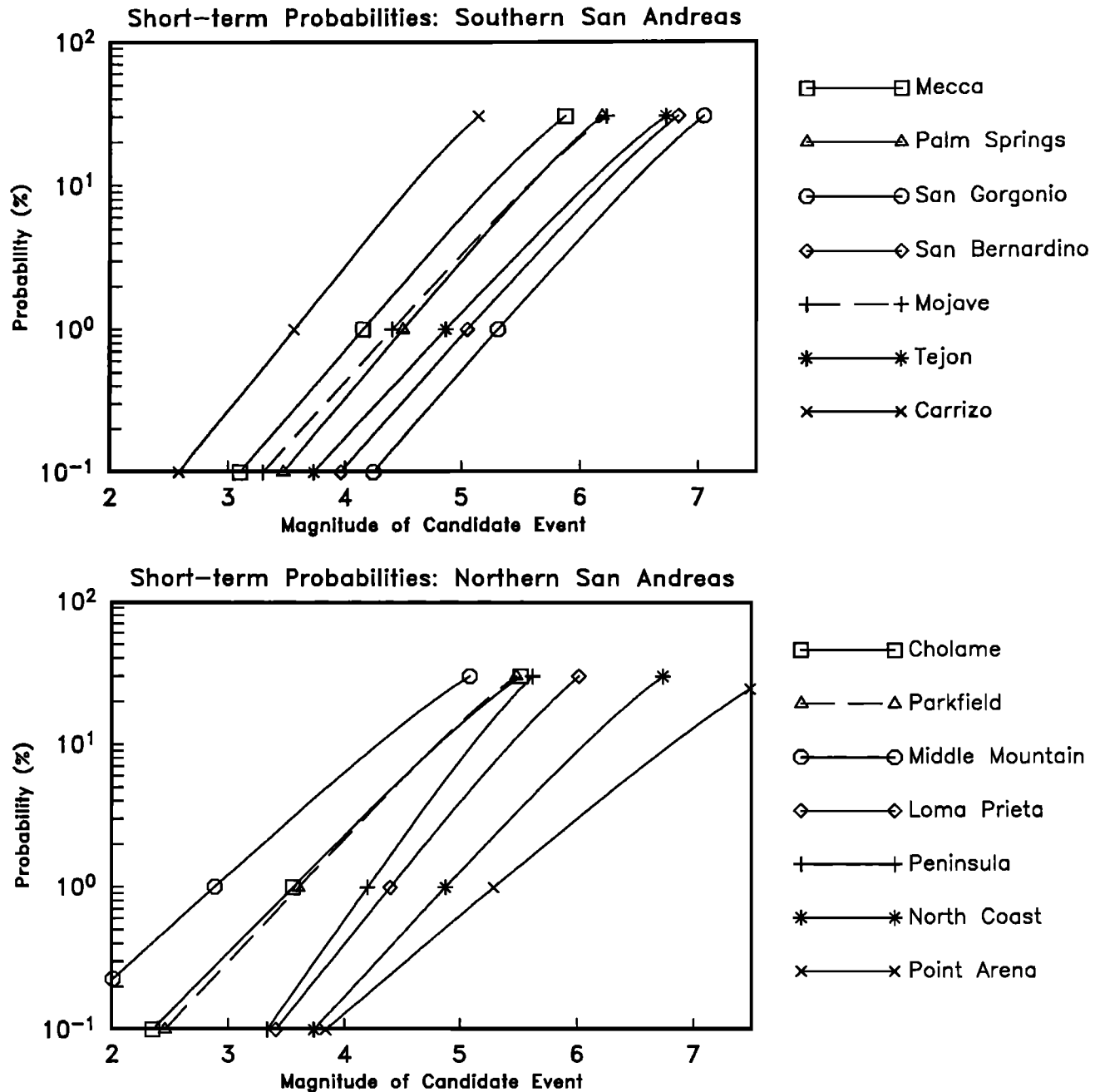
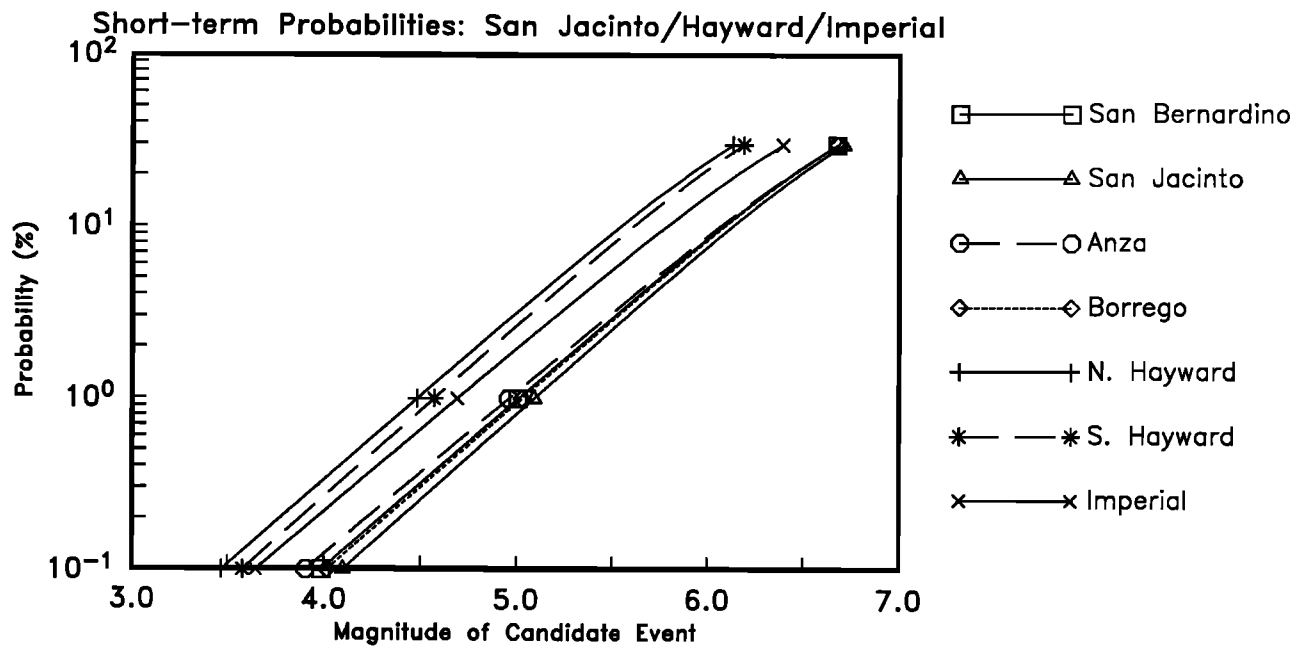


Fig. 7. The probability that an earthquake of given magnitude is a foreshock to a characteristic mainshock, plotted against magnitude for each fault segment listed in Table 1. Shown are results for the (a) southern San Andreas, (b) northern San Andreas, and (c) other faults of the San Andreas system.

long-term probability $P(C)$ and the rate of background seismicity. Outside of the Parkfield and "Middle Mountain" segments, which have very high $P(C)$, the highest short-term probabilities are from the Carrizo and Cholame segments; even though the 30-year probability is only 10% on the Carrizo segment, the background seismicity there is almost nonexistent. The San Francisco Peninsula and the San Bernardino Mountain segments both have a 30-year probability of 20%, but the probabilities in the San Gorgonio subregion are much lower than those near San Francisco because of higher background seismicity. At high magnitudes, the lowest probabilities are for the Point Arena segment because of its very low β , which may be a result of catalog incompleteness at low magnitudes.

The possibility of the next Parkfield earthquake triggering a larger earthquake on the Cholame segment has been much discussed. Our procedure gives a magnitude 6 in Cholame a 52% chance of being a foreshock to a characteristic mainshock on that segment; but this result comes from the low background rate for the Cholame segment itself. Since this rate predicts a magnitude 6 shock every 1400 years, not every 22 years as at Parkfield, this high probability does not apply to a possible Parkfield trigger. We can, however, use (3) of our zero-dimensional model to roughly estimate the probability that a Parkfield earthquake will be a foreshock to a larger earthquake at Cholame. The WGCEP-88 probability of a Cholame earthquake is 30% in 30 years, while the background rate for



Parkfield mainshocks (and hence candidate Cholame foreshocks) is one every 21.7 years. To determine the short-term probability with (3) we need to assume a value for $P(F|C)$, the rate at which Cholame mainshocks have Parkfield mainshocks as foreshocks. If we assume that, as for an average magnitude 7 earthquake, 15% of Cholame earthquakes are preceded by magnitude 6 foreshocks, then the short-term probability of a Cholame earthquake after a Parkfield earthquake is 3%. At the other extreme, if we assume that 50% of Cholame earthquakes are preceded by Parkfield earthquakes (the only type of foreshock it has is Parkfield earthquakes), the probability becomes 10%.

As discussed in section 2.2 above, the rate of false alarms depends on the background probability for the characteristic earthquake. A cumulative false-alarm rate for the whole San Andreas fault is thus dominated by the contribution from Parkfield, for which a 10% probability level occurs every 8.4 years. By comparison, for the Coachella Valley segment of the San Andreas fault the false alarm rate for a 10% probability is once every 63 years. For 0.1% it is once every 5.5 months, but this probability level is only 9 times the background one.

5. DISCUSSION

The procedure developed here can be made more general than has been appropriate for the above application to the San Andreas fault. As discussed in sections 3 and 4, we could include a different dependence of Φ_i on time or make $P(C)$ include information about the most likely epicenters for the mainshock (such as fault jogs or terminations). Another extension would be to set $P(C)$ from an extrapolation of the frequency-magnitude relation; while a violation of the maximum-magnitude model, this would allow application of this technique to many more regions. The greatest flexibility comes from the precurent probability density Φ_{FC} , since we can, as the data warrant, alter this function to include additional data types. For example, there is evidence to suggest that most foreshocks have focal mechanisms similar to that of their

mainshock [Jones and Lindh, 1987]. If that relationship were parameterized, Φ_{FC} and the integration in (15) could include variables describing the difference in focal mechanisms; thus normal- or thrust-faulting earthquakes would be given a lower probability of being a foreshock to a San Andreas mainshock. If any other characteristics are recognized as being more common in foreshocks than background earthquakes (such as number of aftershocks), we can rigorously include this information in our computation of the conditional probabilities. Another direction to go is in improving our estimates for the precurent probability beyond the rather simple forms described above. Considerable work has been done in the last few years on how to estimate multivariate density functions, which is precisely the problem at hand [Silverman 1986]. An obvious question is whether the estimated densities differ significantly between regions; if so, this could reflect significant differences in the nucleation and triggering of large earthquakes.

Of course, nothing in the derivations of section 2 is specific to foreshocks; this procedure can be used for any potential earthquake precursor. Equation (6) shows that what is needed is a long-term mainshock probability $P(C)$, a rate for background events $P(B)$, and a precurent probability Φ_{FC} , which would in many cases just be the fraction of mainshocks with precursors. At present, these data are not available for any precursor but foreshocks. For instance, the background rate of creep events can be determined for some sections of the San Andreas fault system, but we have almost no data on the fraction of mainshocks preceded by such events.

There have been a number of earlier papers on estimating the probabilities of earthquakes in the presence of precursors [Kagan and Knopoff 1977; Vere-Jones 1978; Guagenti and Scirocco 1980; Aki 1981; Anderson 1982; Grandori et al. 1984]. Most of these take a slightly different definition of events from the ones we have used. Rather than distinguishing between background events (independent of large earthquakes) and precursors (always followed by a large earthquake), these papers assume that all possible precursors fall into one class of events, with some probability of a possible precursor not being followed

by an earthquake. (For example, Anderson [1982] computes the probabilities of a precursor being useful or useless). For seismicity, a division into background and precursory events appears to be a better approximation to the likely physics. Most of these papers also deal with the case (not discussed here) of how possible multiple precursors could increase the conditional probability above that for a single precursor. The discussion above suggests that this will usually be a moot point, since only rarely do we have the information needed to estimate the conditional probabilities. With the exception of the work of Kagan and Knopoff [1987] and (in part) Anderson [1982], there does not seem to have been much consideration of any multidimensional cases of the kind described in section 3. The Kagan and Knopoff study is closest to the approach presented here, though the functional form employed by them is derived from a fracture mechanics model, whereas ours is more purely empirical. The models also differ considerably in their specification of long-term probability. In the Kagan and Knopoff model, this is given by a Poisson rate derived from the frequency-magnitude relation (10), whereas here it can be independent of that. As noted in section 2, such independence appears to be a more satisfactory representation of the seismicity of an active fault zone.

6. CONCLUSIONS

We have shown that the probability that an earthquake that occurs near a major fault will be a foreshock to the characteristic mainshock depends on the rate of background earthquake activity on that segment, the long-term probability of the mainshock, and the rate at which the mainshocks are preceded by foreshocks, which we call the precurrent probability. Assuming certain reasonable forms for the density function of this probability (as a function of time, location, and magnitude) we have found an expression for the short-term probability that an earthquake is a foreshock, and applied it to the faults of the San Andreas system. Because the rate of foreshocks before mainshocks is assumed to be the same for all segments, the differences in short-term probabilities between segments arise from differences in background rate of seismicity and in long-term probabilities. The background rates are more variable between regions and lead to larger variations in short-term probabilities. For the San Andreas fault the two extremes are the nearly aseismic Carrizo Plain, where a 1% probability for a characteristic earthquake would be found for a magnitude 3.6 candidate event, and the highly seismic San Geronio region, where it would take a magnitude 5.3 to reach this level.

Acknowledgements. We thank the members of the Working Group on Short-term Earthquake Alerts for the Southern San Andreas Fault, especially Brad Hager and Dave Jackson, for raising some of the issues that led to this paper. We have benefited greatly from reviews by Andy Michael, Dave Jackson (again), Al Lindh, and especially Mark Mathews. We also thank Paul Reasenberg for comments and for providing the declustered CALNET data. Preparation of this paper was in part supported by U.S. Geological Survey grant 14-08-0001-G1763.

REFERENCES

- Aki, K., A probabilistic synthesis of precursory phenomena, in *Earthquake Prediction: An International Review*, edited by P. Richards, pp. 566–574, American Geophysical Union, Washington D. C., 1981.
- Anderson, J. G., Revised estimates for the probabilities of earthquakes following the observation of unreliable precursors, *Bull. Seismol. Soc. Amer.*, 72, 879–888, 1982.
- Bakun, W. H., K. S. Breckenridge, J. Bredehoeft, R. O. Burford, W. L. Ellsworth, M. J. S. Johnston, L. Jones, A. G. Lindh, C. Mortenson, R. J. Mueller, C. M. Poley, E. Roeleffs, S. Schulz, P. Segall, and W. Thatcher, Parkfield, California, earthquake prediction scenarios and response plans, *U. S. Geol. Surv. Open-file Rep.* 86-365, 1987.
- Bakun, W. H., G. C. P. King, and R. S. Cockerham, Seismic slip, aseismic slip, and the mechanics of repeating earthquakes on the Calaveras fault, California, in *Earthquake Source Mechanics (Geophysical Monograph 37)*, edited by C. Scholz, pp. 195–207, American Geophysical Union, Washington D. C., 1986.
- Brune, J. N., Implications of earthquake triggering and rupture propagation for earthquake prediction based on premonitory phenomena, *J. Geophys. Res.*, 84, 2195–2197, 1979.
- Davis, P. M., D. D. Jackson, and Y. Kagan, The longer it has been since the last earthquake, the longer the expected time till the next?, *Bull. Seismol. Soc. Amer.*, 79, 1439–1456, 1989.
- Given, D. D., L. K. Hutton, L. Stach, and L. M. Jones, The Southern California Network Bulletin, January–June, 1987, *U.S. Geol. Surv. Open-file Rep.* 88-408, 1988.
- Goltz, J., The Parkfield and San Diego earthquake predictions: a chronology, *Special Report by the Southern California Earthquake Preparedness Project*, Los Angeles, Calif., 1985.
- Grandori, G., E. Guagenti, and F. Perotti, Some observations on the probabilistic interpretation of short-term earthquake precursors, *Earthq. Eng. Struct. Dynam.*, 12, 749–760, 1984.
- Guagenti, E. G. and F. Scirocco, A discussion of seismic risk including precursors, *Bull. Seismol. Soc. Amer.*, 70, 2245–2251, 1980.
- Jones, L. M., Foreshocks (1966–1980) in the San Andreas System, California, *Bull. Seismol. Soc. Amer.*, 74, 1361–1380, 1984.
- Jones, L. M., Foreshocks and time-dependent earthquake hazard assessment in southern California, *Bull. Seismol. Soc. Amer.*, 75, 1669–1680, 1985.
- Jones, L. M. and A. G. Lindh, Foreshocks and time-dependent conditional probabilities of damaging earthquakes on major faults in California, *Seismol. Res. Letters*, 58, 21, 1987.
- Jones, L. M. and P. Molnar, Some characteristics of foreshocks and their possible relationship to earthquake prediction and premonitory slip on faults, *J. Geophys. Res.*, 84, 3596–3608, 1979.
- Kagan, Y. and L. Knopoff, Earthquake risk prediction as a stochastic process, *Phys. Earth Planet. Int.*, 14, 97–108, 1977.
- Kagan, Y. and L. Knopoff, Statistical short-term earthquake prediction, *Science*, 236, 1563–1567, 1987.
- King, G. C. P. and J. Nábelek, Role of fault bends in the initiation and termination of rupture, *Science*, 228, 984–987, 1985.
- Knott, C. G., *The Physics of Earthquake Phenomena*, 278 pp., Clarendon Press, Oxford, 1908.
- Lindh, A. G., Preliminary assessment of long-term probabilities for large earthquakes along selected segments of the San Andreas fault system in California, *U.S. Geol. Surv. Open-File Rep.* 83-63, pp. 1–15, 1983.
- Nishenko, S. P. and R. Buland, A generic recurrence interval distribution for earthquake forecasting, *Bull. Seismol. Soc. Amer.*, 77, 1382–1399, 1987.
- Reasenberg, P., Second-order moment of Central California seismicity, 1969–1982, *J. Geophys. Res.*, 90, 5479–5496, 1985.
- Reasenberg, P. A. and L. M. Jones, Earthquake hazard after a mainshock in California, *Science*, 243, 1173–1176, 1989.
- Silverman, B., *Density Estimation*, 175 pp., Chapman and Hall, London, 1986.
- Sykes, L. and S. P. Nishenko, Probabilities of occurrence of large plate-rupturing earthquakes for the San Andreas, San Jacinto, and Imperial faults, California 1983–2003, *J. Geophys. Res.*, 89, 5905–5927, 1984.
- Vere-Jones, D., Earthquake prediction—a statistician's view, *J. Phys. Earth*, 26, 129–146, 1978.
- Wesnousky, S. G., C. H. Scholz, K. Shimazaki, and T. Matsuda, Earthquake frequency distribution and faulting mechanics, *J. Geophys. Res.*, 88, 9331–9340, 1983.
- Working Group on California Earthquake Probabilities, Probabilities of large earthquakes occurring in California on the San Andreas fault, *U.S. Geol. Surv. Open-File Rep.* 88-398, 1988.
- D. C. Agnew, Institute of Geophysics and Planetary Physics, University of California, San Diego, La Jolla, CA 92093-0225.
- L. M. Jones, U.S. Geological Survey, 525 S. Wilson Avenue, Pasadena, CA 91106

(Received August 17, 1990;
revised January 14, 1991;
accepted January 18, 1991.)